

DETERMINING STRUCTURAL SIMILARITY IN SEMI-STRUCTURED DOCUMENTS

ABSTRACT

5

Documents are represented based on their structure, which arises from the relationship between various elements in the document. After representing documents based on their structure in vector form, a method of measuring similarity between vectors is used to obtain the measure of structural similarity between two given documents.

10